RAW SEQUENCE LISTING
ERROR REPORT
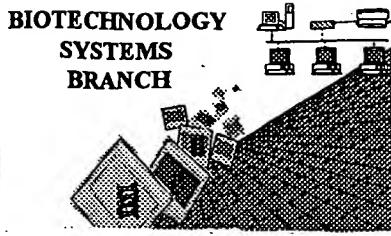
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/941,193_
Source: _1636_
Date Processed by STIC: _3/1/2002_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: **patin21help@uspto.gov** or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: **patin3help@uspto.gov** or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
   http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (**<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
    U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
       Or
    U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/941,193

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino
Numbering

The numbering under each 5ᵗʰ amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length

Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) ____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 ____ Skipped Sequences
(OLD RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences
(NEW RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 ____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 ____ Use of <220>

Sequence(s) ____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

1636

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/941,193

DATE: 03/01/2002
TIME: 14:45:45

Input Set : A:\Seq-sub.app
Output Set: N:\CRF3\03012002\I941193.raw

Does Not Comply
Corrected Diskette Needed

SEQUENCE LISTING
4 (1) GENERAL INFORMATION:
6   (i) APPLICANT: BROW, MARY ANN D.
7   LYAMICHEV, VICTOR I.
8   OLIVE, DAVID M.
10   (ii) TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
11   PATHOGENS
13   (iii) NUMBER OF SEQUENCES: 165
15   (iv) CORRESPONDENCE ADDRESS:
16   (A) ADDRESSEE: MEDLEN & CARROLL
17   (B) STREET: 220 MONTGOMERY STREET, SUITE 2200
18   (C) CITY: SAN FRANCISCO
19   (D) STATE: CALIFORNIA
20   (E) COUNTRY: UNITED STATES OF AMERICA
21   (F) ZIP: 94104
23   (v) COMPUTER READABLE FORM:
24   (A) MEDIUM TYPE: Floppy disk
25   (B) COMPUTER: IBM PC compatible
26   (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27   (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
29   (vi) CURRENT APPLICATION DATA:
30   (A) APPLICATION NUMBER: US/09/941,193
31   (B) FILING DATE: 28-Aug-2001
32   (C) CLASSIFICATION:
34 (viii) ATTORNEY/AGENT INFORMATION:
35   (A) NAME: CARROLL, PETER G.
36   (B) REGISTRATION NUMBER: 32,837
37   (C) REFERENCE/DOCKET NUMBER: FORS-01756
39   (ix) TELECOMMUNICATION INFORMATION:
40   (A) TELEPHONE: (415) 705-8410
41   (B) TELEFAX: (415) 397-8338

ERRORED SEQUENCES

44 (2) INFORMATION FOR SEQ ID NO: 1:
46   (i) SEQUENCE CHARACTERISTICS:
47   (A) LENGTH: 2506 base pairs
48   (B) TYPE: nucleic acid
49   (C) STRANDEDNESS: double
50   (D) TOPOLOGY: linear
52   (ii) MOLECULE TYPE: DNA (genomic)
56   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING   DATE: 03/01/2002
PATENT APPLICATION: US/09/941,193   TIME: 14:45:45

        Input Set : A:\Seq-sub.app
        Output Set: N:\CRF3\03012002\I941193.raw

```
E--> 58 ATGAGGGGGA TGCTGCCCCT CTTTGAGCCC AAGGGCCGGG TCCTCCTGGT
     59 GGACGGCCAC      60
E--> 61 CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG
     62 GGGGGAGCCG     120
E--> 64 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTCAAGGA
     65 GGACGGGGAC     180
E--> 67 GCGGTGATCG TGGTCTTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC
     68 CTACGGGGGG     240
E--> 70 TACAAGGCGG GCCGGGCCCC CACGCCGGAG GACTTTCCCC GGCAACTCGC
     71 CCTCATCAAG     300
E--> 73 GAGCTGGTGG ACCTCCTGGG GCTGGCGCGC CTCGAGGTCC CGGGCTACGA
     74 GGCGGACGAC     360
E--> 76 GTCCTGGCCA GCCTGGCCAA GAAGGCGGAA AAGGAGGGCT ACGAGGTCCG
     77 CATCCTCACC     420
E--> 79 GCCGACAAAG ACCTTTACCA GCTCCTTTCC GACCGCATCC ACGTCCTCCA
     80 CCCCGAGGGG     480
E--> 82 TACCTCATCA CCCCGGCCTG GCTTTGGGAA AAGTACGGCC TGAGGCCCGA
     83 CCAGTGGGCC     540
E--> 85 GACTACCGGG CCCTGACCGG GGACGAGTCC GACAACCTTC CCGGGGTCAA
     86 GGGCATCGGG     600
E--> 88 GAGAAGACGG CGAGGAAGCT TCTGGAGGAG TGGGGGAGCC TGGAAGCCCT
     89 CCTCAAGAAC     660
E--> 91 CTGGACCGGC TGAAGCCCGC CATCCGGGAG AAGATCCTGG CCCACATGGA
     92 CGATCTGAAG     720
E--> 94 CTCTCCTGGG ACCTGGCCAA GGTGCGCACC GACCTGCCCC TGGAGGTGGA
     95 CTTCGCCAAA     780
E--> 97 AGGCGGGAGC CCGACCGGGA GAGGCTTAGG GCCTTTCTGG AGAGGCTTGA
     98 GTTTGGCAGC     840
E--> 100 CTCCTCCACG AGTTCGGCCT TCTGGAAAGC CCCAAGGCCC TGGAGGAGGC
     101 CCCCTGGCCC     900
E--> 103 CCGCCGGAAG GGGCCTTCGT GGGCTTTGTG CTTTCCCGCA AGGAGCCCAT
     104 GTGGGCCGAT     960
E--> 106 CTTCTGGCCC TGGCCGCCGC CAGGGGGGGC CGGGTCCACC GGGCCCCCGA
     107 GCCTTATAAA    1020
E--> 109 GCCCTCAGGG ACCTGAAGGA GGCGCGGGGG CTTCTCGCCA AAGACCTGAG
     110 CGTTCTGGCC    1080
E--> 112 CTGAGGGAAG GCCTTGGCCT CCCGCCCGGC GACGACCCCA TGCTCCTCGC
     113 CTACCTCCTG    1140
E--> 115 GACCCTTCCA ACACCACCCC CGAGGGGGTG GCCCGGCGCT ACGGCGGGGA
     116 GTGGACGGAG    1200
E--> 118 GAGGCGGGGG AGCGGGCCGC CCTTTCCGAG AGGCTCTTCG CCAACCTGTG
     119 GGGGAGGCTT    1260
E--> 121 GAGGGGGAGG AGAGGCTCCT TTGGCTTTAC CGGGAGGTGG AGAGGCCCCT
     122 TTCCGCTGTC    1320
E--> 124 CTGGCCCACA TGGAGGCCAC GGGGGTGCGC CTGGACGTGG CCTATCTCAG
     125 GGCCTTGTCC    1380
E--> 127 CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC GAGGCCGAGG TCTTCCGCCT
     128 GGCCGGCCAC    1440
E--> 130 CCCTTCAACC TCAACTCCCG GGACCAGCTG GAAAGGGTCC TCTTTGACGA
```

global format error

see item 1 on Error Summary Sheet

Due to size of Sequence Listing, these pages only, is shown as a sample of global error.

```
                RAW SEQUENCE LISTING              DATE: 03/01/2002
                PATENT APPLICATION:  US/09/941,193   TIME: 14:45:45

                Input Set : A:\Seq-sub.app
                Output Set: N:\CRF3\03012002\I941193.raw

       131 GCTAGGGCTT     1500
E--> 133 CCCGCCATCG GCAAGACGGA GAAGACCGGC AAGCGCTCCA CCAGCGCCGC
       134 CGTCCTGGAG     1560
E--> 136 GCCCTCCGCG AGGCCCACCC CATCGTGGAG AAGATCCTGC AGTACCGGGA
       137 GCTCACCAAG     1620
E--> 139 CTGAAGAGCA CCTACATTGA CCCCTTGCCG GACCTCATCC ACCCCAGGAC
       140 GGGCCGCCTC     1680
E--> 142 CACACCCGCT TCAACCAGAC GGCCACGGCC ACGGGCAGGC TAAGTAGCTC
       143 CGATCCCAAC     1740
E--> 145 CTCCAGAACA TCCCCGTCCG CACCCCGCTT GGGCAGAGGA TCCGCCGGGC
       146 CTTCATCGCC     1800
E--> 148 GAGGAGGGGT GGCTATTGGT GGCCCTGGAC TATAGCCAGA TAGAGCTCAG
       149 GGTGCTGGCC     1860
E--> 151 CACCTCTCCG GCGACGAGAA CCTGATCCGG GTCTTCCAGG AGGGGCGGGA
       152 CATCCACACG     1920
E--> 154 GAGACCGCCA GCTGGATGTT CGGCGTCCCC CGGGAGGCCG TGGACCCCCT
       155 GATGCGCCGG     1980
E--> 157 GCGGCCAAGA CCATCAACTT CGGGGTCCTC TACGGCATGT CGGCCCACCG
       158 CCTCTCCCAG     2040
E--> 160 GAGCTAGCCA TCCCTTACGA GGAGGCCCAG GCCTTCATTG AGCGCTACTT
       161 TCAGAGCTTC     2100
E--> 163 CCCAAGGTGC GGGCCTGGAT TGAGAAGACC CTGGAGGAGG GCAGGAGGCG
       164 GGGGTACGTG     2160
E--> 166 GAGACCCTCT TCGGCCGCCG CCGCTACGTG CCAGACCTAG AGGCCCGGGT
       167 GAAGAGCGTG     2220
E--> 169 CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC ATGCCCGTCC AGGGCACCGC
       170 CGCCGACCTC     2280
E--> 172 ATGAAGCTGG CTATGGTGAA GCTCTTCCCC AGGCTGGAGG AAATGGGGGC
       173 CAGGATGCTC     2340
E--> 175 CTTCAGGTCC ACGACGAGCT GGTCCTCGAG GCCCCAAAAG AGAGGGCGGA
       176 GGCCGTGGCC     2400
E--> 178 CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG TATCCCCTGG CCGTGCCCCT
       179 GGAGGTGGAG     2460
E--> 181 GTGGGGATAG GGGAGGACTG GCTCTCCGCC AAGGAGTGAT ACCACC
W--> 182 2506
       184 (2) INFORMATION FOR SEQ ID NO: 2:
       186     (i) SEQUENCE CHARACTERISTICS:
       187         (A) LENGTH: 2496 base pairs
       188         (B) TYPE: nucleic acid
       189         (C) STRANDEDNESS: double
       190         (D) TOPOLOGY: linear
       192     (ii) MOLECULE TYPE: DNA (genomic)
       196     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
E--> 198 ATGGCGATGC TTCCCCTCTT TGAGCCCAAA GGCCGCGTGC TCCTGGTGGA
       199 CGGCCACCAC       60
E--> 201 CTGGCCTACC GCACCTTCTT TGCCCTCAAG GGCCTCACCA CCAGCCGCGG
       202 CGAACCCGTT      120
E--> 204 CAGGCGGTCT ACGGCTTCGC CAAAAGCCTC CTCAAGGCCC TGAAGGAGGA
       205 CGGGGACGTG     180
```

see item 1

see item 1